Development of machine learning models to predict ambulation outcomes after spinal metastasis surgery

Piya Chavalparit
Mahidol University
Sirichai Wilartratsami
Mahidol University
Borriwat Santipas
Mahidol University
Piyalitt Ittichaiwong
Mahidol University
Panya Luksanapruksa (Cutecarg@yahoo.com)
Mahidol University

Article

Keywords: Development, machine learning models, predict, ambulation outcomes, surgery, spinal metastasis

Posted Date: April 4th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1436840/v1

Abstract

Background: Postoperative ambulation status after spinal metastasis surgery is currently difficult to predict. Improved ability to predict this important postoperative outcome would improve management decision-making and help in determining realistic goals of treatment. Accordingly, this study set forth to develop machine learning models to predict ambulation outcomes after surgery for spinal metastasis.

Methods: This retrospective study included patients who underwent spinal metastasis at a university-based medical center in Thailand during January 2009-November 2021. Collected data included preoperative parameters, and ambulatory status at 90 and 180 days after surgery. Seven machine learning algorithms, including decision tree, random forest, XGBoost, logistic regression, support vector machine (SVM), neural network, and stochastic gradient descent (SGD), were developed to predict ambulatory status at 90-days and 180-days postoperation. Model performance was evaluated using the area under the receiver operating characteristic curve (AUC) and F1-score.

Results: A total of 178 patients were enrolled. The number of patients classified as ambulatory at 90-days and 180-days after surgery was 150 (84.3%) and 145 (81.5%) respectively. The XGBoost algorithm was found to most accurately predict 180-day ambulatory outcome (AUC: 0.92, F1-score: 0.92), and the random forest algorithm was shown to most accurately predict 90-day ambulatory outcome (AUC: 0.95, F1-score: 0.94).

Conclusion: Machine learning algorithms were shown to be effective for predicting ambulatory status after surgery for spinal metastasis. The XGBoost and random forest algorithms best predicted postoperative ambulatory status at 180-days and 90-days after spinal metastasis surgery, respectively.

Introduction

The incidence of spinal metastasis is increasing as evidenced by recent studies that reported that 5%-10% of all cancer patients develop spinal metastasis [1–3]. To determine treatment options, several factors must be considered, such as disease factors, patient factors, and patient expectations. Taking all of these factors into account, it is important to establish clear and realistic goals of treatment [4,5].

Treatments for spinal metastasis have rapidly improved to maximize survival and clinical outcomes [6]. However, despite advancements in treatment, some patients continue to have poor clinical outcomes and are unable to ambulate after spinal metastasis surgery [7–10]. Previous study proposed models for predicting ambulatory ability after spinal metastasis surgery that were developed using conventional statistical methods, but those models yielded only fair to moderate performance [11].

To yield improved benefit from vast amounts of exponentially generated data, artificial intelligence (AI) and machine learning (ML) were recently employed to develop new tools to improve spine treatment and research [12,13]. A number of applications using ML in the field of spine surgery were reported with promising results that outperformed conventional statistical methods [14–17].

Postoperative ambulation status after spinal metastasis surgery is difficult to predict, and improved ability to predict this important postoperative outcome would improve management decision-making and help in determining clear and realistic goals of treatment. Accordingly, the aim of this study was to develop machine learning algorithms to predict ambulation outcomes after surgery for spinal metastasis.

Materials And Method

Guidelines

This study followed the Transparent Reporting of a Multivariable Prediction Models for Individual Prognosis or Diagnosis (TRIPOD) guidelines, and the Guidelines for Developing and Reporting Machine Learning Models in Biomedical Research. All methods were performed in accordance with the relevant guidelines, regulations and the declaration of Helsinki. The study protocol was approved by the Siriraj Institutional Review Board (COA no. 978/2021, 937/2564[IRB1]).

Patient selection

Consecutive patients who underwent surgery for spinal metastasis at a university-based medical center in Thailand during January 2009 to November 2021 were retrospectively enrolled. The inclusion criteria were (i) diagnosis of spinal metastasis, (ii) 18 years of age or older, and (iii) underwent surgery for cervical, thoracic, lumbar, and/or sacral metastasis/metastases. Patients who expired before 180 days after surgery or who had no record of their ambulatory status at 180 days after surgery were excluded. Written informed consent was waived by the Siriraj Institutional Review Board due to the retrospective nature of this study.

Variables

Preoperative parameters were collected via retrospective chart review. Factors that were previously reported to be significantly associated with ambulatory status after spinal metastasis surgery [15,18–22] were collected, including age, gender, body mass index (BMI) (kg/m^2), smoking status, American Society of Anesthesiologists (ASA) classification, presence of myelopathy before surgery, duration of neurological deficit, Frankel grading, level of spinal compression, level of spinal metastasis, comorbidities, extraspinal bone metastasis, visceral metastasis, preoperative treatment (chemotherapy, radiotherapy, targeted therapy), primary tumor origin, serum calcium level, albumin level, creatinine, surgical procedure, level of surgery, and preoperative ambulatory status. Primary tumor histology was also included to more fully and clearly describe the primary tumor.

Outcomes

Previous study reported that functional recovery reaches the plateau phase at 6 months after spinal metastasis surgery [23]. We, therefore, selected ambulatory status at 180 days after surgery as the primary outcome, and ambulatory status at 90 days after surgery as the secondary outcome. Ambulatory status as ‘ambulator’ was defined as patients who can walk without assistance (with or without gait aid).

Patients who were unable to walk without assistance were classified as ‘non-ambulators’. Predictors and outcomes were independently reviewed by two different orthopedic surgeons for blind assessment.

Missing data

Patients with no primary or secondary outcome data were excluded. In cases with missing preoperative data, multiple imputations with chained equations was used.

Prediction models

The ML models included in this study were used in previous study to evaluate survival among patients with metastatic disease [24]. The performance of all included ML models was compared to identify the best performing model for both the primary and secondary outcome.

Seven ML models included decision tree, random forest, XGBoost, logistic regression, support vector machine (SVM), neural network, and stochastic gradient descent (SGD). All models were created with Python 3.9 using Scikit-learn library v.1.0.1 under open source simplified BSD license [25].

The dataset was divided into the training set and the testing set at an 80:20 ratio. Model training was conducted using the training set with performance validation by 5-fold cross-validation. A class weighting strategy was also used to ensure that the trained model would take each class into equal account despite class imbalance. Model performance was evaluated using the testing data set, and by evaluating and comparing the area under the receiver operating characteristic curve (AUC) and F1-score among the 7 models. An AUC of 0.7–0.8 indicates fair performance, and an AUC of greater than 0.8 indicates good performance. The F1-score, which is calculated using the precision and recall parameters, has a maximum possible value of 1.0, which indicates perfect performance.

Results

Although a total of 405 patients with spinal metastasis met the inclusion criteria, only 245 of those patients were still alive at 180 days after spinal metastasis surgery. Patients without recorded outcome data were also excluded. In the end, 178 participants were enrolled in this study, including 79 men (44.4%) and 99 women (55.6%). The mean age of all patients was 57.0 ± 11.6 years. The number of patients classified as ambulatory 90 days and 180 days after spinal metastasis surgery was 150 (84.3%) and 145 (81.5%), respectively.

Types of missing data included BMI in 9 (5.0%) patients, serum calcium level in 22 (12.0%) patients, serum creatinine level in 4 (2.2%) patients, surgical procedure in 1 (0.5%) patient, and level of surgery in 3 (1.6%) patients. Multiple imputations with chained equation was used to impute the missing data. Baseline characteristics compared between the ambulatory and non-ambulatory at 180 days groups are shown in Table 1.

The factors selected by the XGBoost algorithm that significantly predict 180-day ambulatory outcome were presence of symptomatic spinal compression at thoracic level (27%), histological finding of primary tumor (14%), serum creatinine level (13%), the primary source of tumor (12%), and cervical spinal metastasis (12%). The factors selected by the random forest algorithm that significantly predict 90-day ambulatory outcome were preoperative ambulatory status (11%), symptomatic spinal compression at thoracic level (8%), age (8%), serum albumin level (5%), and preoperative neurological status (5%).

Model evaluation for prediction of 180-day ambulatory outcome

Among the 7 models that were evaluated, the XGBoost algorithm was shown to have the best performance for predicting 180-day ambulatory outcome [AUC: 0.92 (Fig. 1), accuracy: 0.89, precision: 0.96, recall: 0.88, and F1-score: 0.92. Data specific to the 180-day prediction performance of all evaluated models are given in Table 2.

Model evaluation for prediction of 90-day ambulatory outcome

Of the seven evaluated models, the random forest algorithm demonstrated the best ability to predict 90-day postoperative ambulatory outcome [AUC: 0.95 (Fig. 2), accuracy: 0.92, precision: 1.00, recall: 0.88, and F1-score: 0.94]. Details relating to the 90-day prediction performance of all models are shown in Table 3.

Discussion

Previous studies reported the benefit of surgery in spinal metastasis relative to regaining ambulatory status, pain-relief [8], quality of life score, and functional outcome score [23]. However, despite promising results from surgery, 3.6–15.3% of patients remained dependent and postoperative complications were as high as 29–34% [7–10]. Consistent with the rates reported from previous studies, 81.5% of spinal metastasis patients in our study were ambulatory at 180 days after surgery.

Factors previously reported to be significantly associated with postoperative clinical outcome were baseline health-related quality of life (HRQOL), preoperative functional status, preoperative neurological function, the interval from symptom to treatment, and the chronology of motor deficit progression [10,19]. Our study demonstrated similar factors for the 90-day outcome, which related to preoperative patient status, including preoperative neurological status and ambulatory status. In contrast, the factors found to be most associated with 180-day postoperative ambulatory outcome were disease-related factors, such as primary source of the tumor, histology, level of compression, and level of metastasis. A possible explanation for this finding may be that after a period of recovery, the effects of surgery may decrease, and the natural course of the disease may become more dominant.

Determining a clear and realistic treatment goal requires accurate information. Previous studies proposed models to predict ambulatory status after spinal metastasis surgery that were developed using conventional statistical methods. Ohashi, *et al*. retrospectively reviewed 82 cases and reported recovery of ambulatory status to be correlated with a duration from the onset of neurological symptoms to gait disability of fewer than 5 days (AUC: 0.72) and a Tokuhashi score of less than 7.5 points (AUC: 0.71) [11]. In the present study, we successfully developed seven ML algorithms and identified the best predictive model for ambulatory status at 180 days (XGBoost) and 90 days (random forest) after surgery with AUC values of 0.92 and 0.95, respectively.

In the 180-day and 90-day study groups, the XGBoost model and random forest model yielded the best results, respectively. In contrast to the SVM and neural network algorithms, the XGBoost and random forest models were originally developed using a decision tree-based model, which is a practical strategy for evaluating relatively small imbalanced-class datasets, such as those used in the present study. This may explain why the XGBoost and random forest models outperformed the other ML models included in this study.

Since most patients in our study were ambulators at 180 days after surgery, this imbalance in data adversely affected ML algorithm development. To remedy this issue, we used a class weighting strategy to optimize the training process, and we included the F1-score for model evaluation. The F1-score provides valuable insights as a metric for examining imbalanced datasets. The XGBoost model, which was shown to best predict ambulatory status at 180 days after surgery, had a very high F1-score of 0.92. Another common problem when developing an ML model is overfitting. To counter the potential of overfitting, we implemented 5-fold cross-validation to continuously monitor model performance during training. Each model was then further evaluated using the testing dataset.

The previously published SORG ML algorithm has been widely adopted for treatment decision-making and to predict survival in spinal metastatic disease patients[17]. However, in addition to the survival rate, postoperative ambulatory status is also a very important factor that needs to be considered. This is the first study to report models that predict 180-day and 90-day ambulatory outcome after spinal metastasis surgery using ML algorithms. Our combined ML model, which allows the user to predict either 180-day or 90-day ambulatory status after surgery, has been deployed as an open access web application that can be found at…

https://share.streamlit.io/orthosiriraj/outcome_post_op_metas_spine/main/main.py

Limitations

This study is limited by its retrospective single-center design. Moreover, our center is a national tertiary referral hospital, which could limit the generalizability of our findings to other care settings. Another potential limitation is that the relatively small amount of included data could limit of performance of ML. To remedy this limitation and continue to improve the performance of our developed algorithms, we will

continue to collect data that will be used to refine the performance of our ML models. Further multicenter study and external validation are needed to shore up the results of this study and to establish the validity of these algorithms for use in real-world clinical practice.

Conclusion

Machine learning algorithms were shown to be effective for predicting ambulatory status after surgery for spinal metastasis. The XGBoost and random forest algorithms best predicted postoperative ambulatory status at 180-days and 90-days after spinal metastasis surgery, respectively. Once externally validated for use in routine clinical practice, these algorithms will improve case management decision-making and help in determining clear and realistic goals of treatment.

Declarations

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Author contributions

PL, SW, PI, and PC designed the study. PC, BS, and PI collected, analyzed the data, and contributed substantially to interpretation of data. PL supervised the project. PC, BS, and PI drafted the article. All authors have read and approved the manuscript.

Conflict of interest declaration

All authors declare no personal or professional conflicts of interest relating to any aspect of this study.

Funding disclosure

This was an unfunded study.

Acknowledgements

The authors gratefully acknowledge Miss Nunnapat Kangkano of the Research Unit of the Department of Orthopaedic Surgery, Faculty of Medicine Siriraj Hospital, Mahidol University for her assistance with statistical analysis, manuscript preparation, and coordination of the journal submission process.

References

1. Klimo Jr., P. & Schmidt, M. H. Surgical Management of Spinal Metastases. The Oncologist **9**, 188–196, doi:https://doi.org/10.1634/theoncologist.9-2-188 (2004).

2. Luksanapruksa, P. *et al.* Epidemiologic Study of Operative Treatment for Spinal Metastasis in Thailand: A Review of National Healthcare Data from 2005 to 2014. J Korean Neurosurg Soc **0**, doi:10.3340/jkns.2020.0330 (2021).
3. Choi, S. H., Koo, J. W., Choe, D. & Kang, C.-N. The Incidence and Management Trends of Metastatic Spinal Tumors in South Korea: A Nationwide Population-based Study. Spine **45**, E856-E863, doi:10.1097/brs.0000000000003445 (2020).
4. Barton, L. B. *et al.* Clinician Experiences in Treatment Decision-Making for Patients with Spinal Metastases: A Qualitative Study. J Bone Joint Surg Am **103**, e1, doi:10.2106/jbjs.20.00334 (2021).
5. Lape, E. C. *et al.* Patient experiences of decision-making in the treatment of spinal metastases: a qualitative study. Spine J **20**, 905–914, doi:10.1016/j.spinee.2019.12.018 (2020).
6. Heary, R. F. & Bono, C. M. Metastatic spinal tumors. Neurosurg Focus **11**, e1, doi:10.3171/foc.2001.11.6.2 (2001).
7. Alamanda, V. K., Robinson, M. M., Kneisl, J. S. & Patt, J. C. Functional and survival outcomes in patients undergoing surgical treatment for metastatic disease of the spine. J Spine Surg **4**, 28–36, doi:10.21037/jss.2018.03.12 (2018).
8. Kim, J. M. *et al.* Clinical outcome of metastatic spinal cord compression treated with surgical excision ± radiation versus radiation therapy alone: a systematic review of literature. Spine (Phila Pa 1976) **37**, 78–84, doi:10.1097/BRS.0b013e318223b9b6 (2012).
9. *The Spine Journal* **16**, 322–328, doi:10.1016/j.spinee.2015.11.005 (2016).
10. Liu, Y. H. *et al.* Prognostic Factors of Ambulatory Status for Patients with Metastatic Spinal Cord Compression: A Systematic Review and Meta-Analysis. World Neurosurg **116**, e278-e290, doi:10.1016/j.wneu.2018.04.188 (2018).
11. Ohashi, M. *et al.* Preoperative prediction for regaining ambulatory ability in paretic non-ambulatory patients with metastatic spinal cord compression. Spinal Cord **55**, 447–453, doi:10.1038/sc.2016.145 (2017).
12. Galbusera, F., Casaroli, G. & Bassani, T. Artificial intelligence and machine learning in spine research. JOR Spine **2**, e1044, doi:10.1002/jsp2.1044 (2019).
13. Rasouli, J. J. *et al.* Artificial Intelligence and Robotics in Spine Surgery. Global Spine J **11**, 556–564, doi:10.1177/2192568220915718 (2021).
14. Merali, Z. G., Witiw, C. D., Badhiwala, J. H., Wilson, J. R. & Fehlings, M. G. Using a machine learning approach to predict outcome after surgery for degenerative cervical myelopathy. PLoS One **14**, e0215133, doi:10.1371/journal.pone.0215133 (2019).
15. Paulino Pereira, N. R. *et al.* The SORG nomogram accurately predicts 3- and 12-months survival for operable spine metastatic disease: External validation. J Surg Oncol **115**, 1019–1027, doi:10.1002/jso.24620 (2017).
16. Ahmed, A. K. *et al.* Predicting survival for metastatic spine disease: a comparison of nine scoring systems. Spine J **18**, 1804–1814, doi:10.1016/j.spinee.2018.03.011 (2018).

17. Karhade, A. V. *et al.* External validation of the SORG 90-day and 1-year machine learning algorithms for survival in spinal metastatic disease. Spine J **20**, 14–21, doi:10.1016/j.spinee.2019.09.003 (2020).
18. Moon, K. Y., Chung, C. K., Jahng, T. A., Kim, H. J. & Kim, C. H. Postoperative survival and ambulatory outcome in metastatic spinal tumors: prognostic factor analysis. J Korean Neurosurg Soc **50**, 216–223, doi:10.3340/jkns.2011.50.3.216 (2011).
19. Feghali, J. *et al.* Predicting postoperative quality-of-life outcomes in patients with metastatic spine disease: who benefits? J Neurosurg Spine, 1–7, doi:10.3171/2020.7.Spine201136 (2020).
20. Luksanapruksa, P. *et al.* Prognostic factors in patients with spinal metastasis: a systematic review and meta-analysis. Spine J **17**, 689–708, doi:10.1016/j.spinee.2016.12.003 (2017).
21. Cheung, Z. B. *et al.* Impact of Obesity on Surgical Outcomes Following Laminectomy for Spinal Metastases. Global Spine Journal **9**, 254–259, doi:10.1177/2192568218780355 (2019).
22. Truong, V. T. *et al.* Surgical Intervention for Patients With Spinal Metastasis From Lung Cancer: A Retrospective Study of 87 Cases. Clin Spine Surg **34**, E133-e140, doi:10.1097/bsd.0000000000001062 (2021).
23. Paulino Pereira, N. R. *et al.* Quality of Life Changes After Surgery for Metastatic Spinal Disease: A Systematic Review and Meta-analysis. Clin Spine Surg, doi:10.1097/bsd.0000000000001213 (2021).
24. Thio, Q. *et al.* Development and Internal Validation of Machine Learning Algorithms for Preoperative Survival Prediction of Extremity Metastatic Disease. Clin Orthop Relat Res **478**, 322–333, doi:10.1097/corr.0000000000000997 (2020).
25. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. **12**, 2825–2830 (2011).

Tables

Table 1. Baseline characteristics of patients who underwent surgery for spinal metastasis compared between the ambulatory and non-ambulatory at 180 days groups

Variables	Missing data (%)	Total	Ambulatory at 180 days (n=145, 81.5%)	Non-ambulatory at 180 days (n=33, 18.5%)
Age (years)	0.0%	178	Mean: 52.2, SD: 11.3	Mean 59.7, SD 9.3
Gender	0.0%	178	Male 61 (42.0%), Female 84 (58.0%)	Male 18 (54.5%), Female 15 (45.5%)
Body mass index (kg/m^2)	5.1%	169	Mean: 22.05, SD: 4.52	Mean 22.80, SD 3.37
Current smoker	0.0%	178	Yes 14 (9.7%), No 131 (90.3%)	Yes 5 (15.2%), No 28 (84.8%)
ASA classification	0.0%	178	Class II 117 (80.7%), Class III 28 (19.3%)	Class II 28 (84.8%), Class III 5 (15.2%)
Comorbidities	0.0%	178	Neurogenic bladder 25 (17.2%), Pneumonia 2 (1.4%), Stroke 1 (0.7%), Myocardial infarction 1 (0.7%), Delirium 2 (1.4%)	Neurogenic bladder 15 (45.5%), Pneumonia 1 (3.0%), Myocardial infarction 1 (3.0%), Delirium 1 (3.0%)
Presence of myelopathy	0.0%	178	Yes 47 (32.4%), No 98 (67.6%)	Yes 15 (45.5%), No 18 (54.5%)
Duration of neurological deficit (days)	0.0%	178	Mean: 15.2, SD: 22.3	Mean: 17.3, SD: 31.4
Frankel grading	0.0%	178	A 1 (0.7%), B 9 (6.2%), C 43 (29.7%), D 72 (49.7%), E 20 (13.8%)	A 0 (0.0%), B 11 (33.3%), C 14 (42.4%), D 5 (15.2%), E 3 (9.1%)
Symptomatic spinal compression	0.0%	178	Cervical 18 (12.4%), Thoracic 37 (25.5%), Lumbar 1 (0.7%), Sacrum 1 (0.7%)	Cervical 2 (6.1%), Thoracic 26 (78.8%), Lumbar 5 (15.2%)
Level of metastasis	0.0%	178	Cervical 36 (24.8%), Thoracic 83 (57.2%), Lumbar 83 (57.2%), Sacrum 18 (12.4%)	Cervical 6 (18.2%), Thoracic 29 (87.9%), Lumbar 8 (24.2%), Sacrum 1 (3.0%)
Extraspinal bone metastasis	0.0%	178	Yes 71 (49.0%), No 74 (51.0%)	Yes 13 (39.4%), No 20 (60.6%)
Visceral metastasis	0.0%	178	Yes 35 (24.1%), No 110 (75.9%)	Yes 5 (15.2%), No 28 (84.8%)
Primary tumor source	0.0%	178	Breast 42 (29.0%), Thyroid 10 (6.9%), Kidney 5 (3.4%), Lung 22 (15.2%), Prostate 12 (8.3%), Liver 5 (3.4%), Colorectal 5 (3.4%), Cervix 1 (0.7%), Hematologic 5 (3.4%), Cholangiocarcinoma 4 (2.8%), Nasopharyngeal 5 (3.4%),	Breast 6 (18.2%), Thyroid 2 (6.1%), Kidney 2 (6.1%), Lung 4 (12.1%), Prostate 5 (15.2%), Liver 2 (6.1%), Hematologic 2 (6.1%), Cholangiocarcinoma 2 (6.1%), Nasopharyngeal 2

			Unknown 18 (12.4%), Others 11 (7.6%)	(6.1%), Unknown 5 (15.2%), Others 1 (3.0%)
Primary tumor histology	0.0%	178	Adenocarcinoma 88 (60.7%), Squamous 11 (7.6%), Follicular 6 (4.1%), Small cell 1 (0.7%), Clear cell 1 (0.7%), Unknown 36 (24.8%)	Adenocarcinoma 17 (51.5%), Squamous 3 (9.1%), Follicular 1 (3%), Clear cell 3 (9.1%), Unknown 9 (27.3%)
Preoperative calcium (mg/dL)	12.0%	157	Mean: 9.02, SD: 1.28	Mean: 8.98, SD: 0.67
Preoperative albumin (g/dL)	0.0%	178	Mean: 3.95, SD: 0.53	Mean: 3.78, SD: 0.45
Preoperative creatinine (mg/dL)	1.7%	175	Mean: 0.91. SD: 0.94	Mean: 0.76, SD: 0.54
Preoperative treatment	0.0%	178	Molecular targeting therapy 6 (4.1%), Chemotherapy 53 (36.6%), Radiotherapy 42 (29.0%)	Chemotherapy 9 (27.3%), Radiotherapy 6 (18.2%)
Preoperative ambulatory status	0.0%	178	Ambulator 83 (57.2%), Non-ambulator 62 (42.8%)	Ambulator 5 (15.2%), Non-ambulator 28 (84.8%)

Abbreviations: SD, standard deviation; ASA, American Society of Anesthesiologists

Table 2. Machine learning model performance for predicting 180-day postsurgical ambulatory outcome

Model	AUC	Accuracy	Precision	Recall	F1-score	Calibration loss
XGBoost	0.92	0.89	0.96	0.88	0.92	0.26
Logistic regression	0.84	0.89	0.96	0.88	0.92	0.32
Random forest	0.91	0.89	1.00	0.85	0.92	0.18
SVM	0.82	0.78	0.78	0.96	0.86	0.19
SGD	0.55	0.69	0.73	0.92	0.81	0.33
Neural network	0.67	0.72	0.72	1.00	0.84	0.20
Decision tree	0.76	0.83	0.88	0.88	0.88	0.20

Abbreviations: AUC, area under the receiver operating characteristic curve; SVM, support vector machine; SGD, stochastic gradient descent

Table 3. Machine learning model performance for predicting 90-day postsurgical ambulatory outcome

Model	AUC	Accuracy	Precision	Recall	F1-score	Calibration loss
Random forest	0.95	0.92	1.00	0.88	0.94	0.20
Logistic regression	0.94	0.86	0.86	0.96	0.91	0.22
XGBoost	0.92	0.86	0.86	0.96	0.91	0.14
Decision tree	0.70	0.72	0.71	1.00	0.83	0.22
Neural network	0.84	0.75	0.75	0.96	0.84	0.21
SVM	0.88	0.83	0.88	0.88	0.88	0.17
SGD	0.79	0.81	0.85	0.88	0.86	0.19

Abbreviations: AUC, area under the receiver operating characteristic curve; SVM, support vector machine; SGD, stochastic gradient descent

Figures

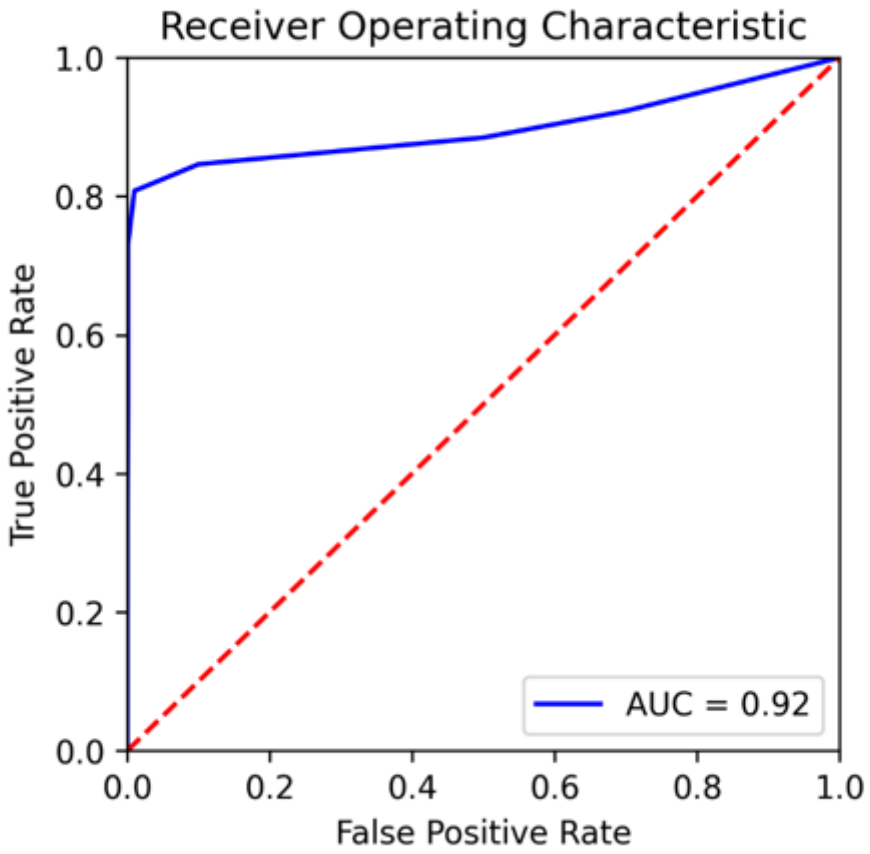

Figure 1

Receiver operating characteristic (ROC) curve for the XGBoost model for predicting 180-day ambulatory status. The area under the ROC curve (AUC) is 0.92.

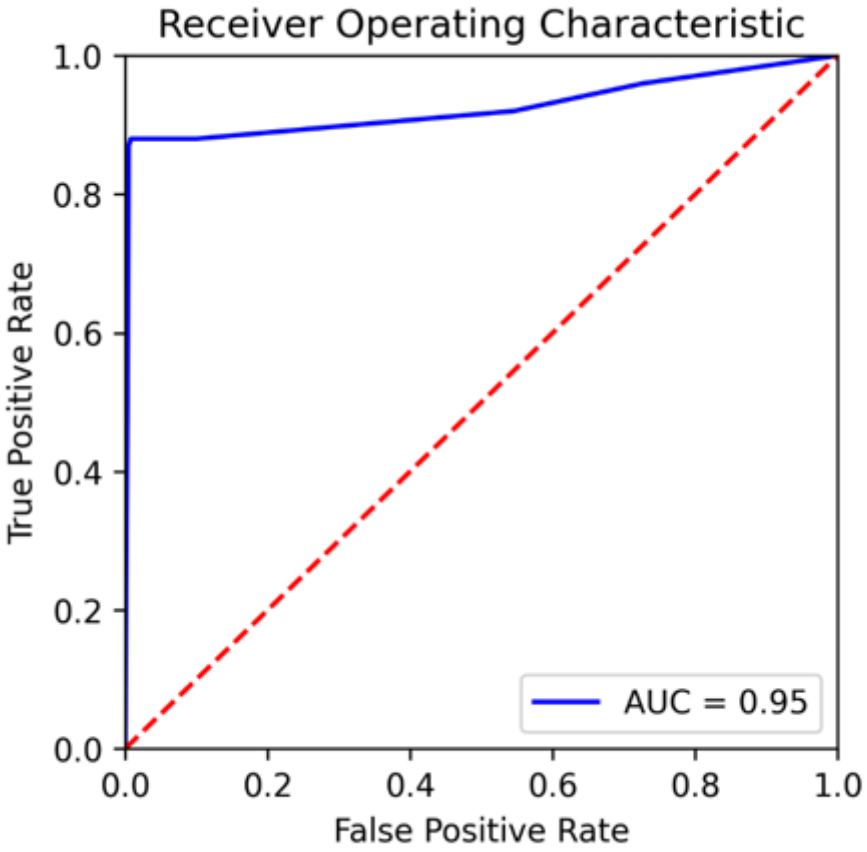

Figure 2

Receiver operating characteristic (ROC) curve for the random forest model for predicting 90-day ambulatory status. The area under the ROC curve (AUC) is 0.95.